# CLASS SIZE, TEACHER PAY, AND STUDENT PERFORMANCE

Presented in Partial Fulfillment of the requirements for

The Master of Education Degree in the

College of Education and Human Service Professions

By

Nan Gao

University of Minnesota Duluth

2012

Committee Signatures:

Chair

Member

Graduate Program Director

## Abstract

Research showed that US education had fallen behind in global competition. To enhance American competitiveness, this study explored two ways to improve education in the US through a cross-country study of education systems in 40 countries. Specifically, this study focused on two important determinants of students' academic performance: class size and teacher quality. This was the first study that used the latest Program for International Student Assessment data to examine the relationship between teacher pay and students' academic performance in a cross-country context. A rigorous empirical analysis showed that while teacher pay had positive, statistically significant and large effects on students' academic performance, class size had no significant effects. Based on this evidence, this study recommended that US government should spend its limited financial resources raising teacher pay. Such a policy would improve the teacher quality and students' academic performance in the US, which would in turn enhance American competitiveness.

TABLE OF CONTENTS

# Chapter 1

## Introduction

On March 20, 2012, US Council on Foreign Relations published a report about US education in the context of global competition. They concluded that there was an "education crisis" in US educational system, which had even evolved into a "national security crisis." To help deal with the "education crisis", this study explored ways to improve education in the US through a cross-country study of education systems in 40 countries. This study focused on two important determinants of students' academic performance: class size and teacher quality.

Regarding class size, conventional wisdom holds that the smaller the classes, the better the school quality, and hence, the greater academic achievements of students. Regarding teacher quality, it depends on many factors. Teacher pay is a particularly important factor that determines teacher quality. This is because to attract and retain the best talents to be teachers, the teacher pay should be set at a proper level. Recently, many researchers (e.g., Lakdawalla, 2006; Eide, Goldhaber and Brewer, 2004) have pointed out teachers are underpaid in the US, so they argue that teacher pay should be raised across the board. However, either reducing class size or increasing teacher pay would require significant financial resources from the government. Under the current budget situation, government financial resources need to be used on the most important determinant of students' academic performance. It is therefore very important to evaluate the effectiveness of two policies to potentially improve students' academic performance in the US: class size reduction, or teacher pay raise.

Class size is an important determinant of student performance, and the conventional wisdom is that smaller classes are associated with better student performance. For example, Halbach et al. (2001) argued that small classes resulted in fewer discipline problems, and teachers could devote more time and individual attention to students. Small classes might also allow more flexibility in instructional strategies. Based on this “conventional wisdom”, in 1996, California implemented the class-size-reduction (CSR) program that reduced K-3 classes throughout the state by roughly 10 students per class, from 30 to 20, at an annual cost that exceeded one billion dollars (Jepsen and Rivkin, 2009). In 2000, the U.S. Congress allocated 1.3 billion dollars for the class size reduction provision of the Elementary and Secondary Education Act (ESEA) (Johnson, 2002). These very costly class size reduction reforms lacked solid empirical evidence to test their effectiveness and made the effect of class size on student performance a highly controversial issue.

In fact, some educational practitioners have already cast doubt on this “conventional wisdom” as well as the policies that have built upon it. They argue that the costs of smaller classes far exceed the benefits. For example, Dr. Eva Moskowitz in ***Washington Post*** (March 3, 2011) vividly describes the high cost of maintaining small classes in New York City:

“In fifth grade, for example, every student gets a laptop and a Kindle with immediate access to an essentially unlimited supply of e-books. Every classroom has a Smart Board, a modem blackboard that is a touch-screen computer with high-speed Internet access. Every teacher has a laptop, video camera, access to a catalogue of lesson plans and videotaped lessons. Outfitting a classroom this way costs about $40,000, or $13,500 amortized over three years. That's how much New York charter schools receive per pupil annually, so we can afford this by just increasing class size by a single student."

To resolve the controversy around the effects of small class sizes and assess the benefits and cost of small class sizes, a systematic analysis was needed to quantify the benefits associated with small class sizes. For such a systematic analysis, this study used a new dataset collected through Program for International Student Assessment (PISA) to examine the effects of class sizes on the academic achievements of students. Such systematic analysis provided the latest evidence on the effects of small class sizes on student performance.

In addition to class size, teacher pay is expected to be an important determinant of student performance. Indeed, Hanushek and Rivkin (2006) showed that teacher quality was the most important input schools contributed to the academic success of their students. Unfortunately, at the current teacher pay level in the US, it has become more and more difficult to attract the best people to become teachers.

For example, Allegretto, Corcoran, and Mishel (2004) examined trends in the relative weekly earnings of elementary and secondary school teachers. In that study they found that the average weekly pay of teachers in 2003 was nearly 14% below that of workers with similar education and work experience. Corcoran, Evans, and Schwab (2004) found that the teacher pay has stagnated in recent years, which has caused a small and declining fraction of the most cognitively skilled graduates who choose to become teachers.

However, there was no consensus among policy makers that teacher pay should be raised in order to attract the best and brightest people to become teachers. For example, on February 1, 2012, Alabama state Senator Shadrack McGill (R) argued that increasing teacher salaries would only lead to less-qualified teachers. In ***Forbes*** magazine on 12/22/2011, Warren Meyer published an article titled "The Teacher Salary Myth—Are Teachers Underpaid?" Mr. Meyer argued that since teachers were paid for 9 months instead of 12 months, and teachers did a lot of work at home, teacher salary was not low at all. Given this controversy, it was important to use the most recent data to conduct a systematic study of the effects of teacher pay on student performance.

**Research Problem**

Class size mattered. First, smaller class sizes resulted in larger educational expenditure. For example, as Krueger (2003, p. 34) argued, "looking across school districts in Texas, variability in the pupil-teacher ratio accounts for two-thirds of the variability in expenditures per student." Second, smaller class sizes were assumed to generate better educational outcomes. For example, Barro and Lee (2001) found that smaller class sizes in primary schools were associated with significantly better student performance. However, Fuchs and Woessmann (2007) and Jurges and Schneider (2004) found no positive relationship between reduced class sizes and student performance. After surveying a large number of papers, Hanushek (1999) concluded that "There is little systematic gain from general reduction in class size'"(p. 33). Similarly, Finn and Petrilli argued, "The conventional wisdom that students do better in smaller classes is flat wrong." (1998, p.382) The problem was clear here: given that smaller classes are expensive to maintain, were they really worth the high costs? In other words, were smaller classes really associated with better student learning outcomes to justify such high costs?

Similarly, teacher compensation was important. But it was controversial whether teachers in the US were underpaid or not. Even though Allegretto, Corcoran, and Mishel (2004) and Corcoran, Evans, and Schwab (2004) found that teachers were underpaid, Senator McGill and ***Forbes*** magazine writer Warren Meyer argued that teachers were not underpaid. The problem was clear here: would higher teacher pay lead to better student performance?

These were important research questions, given the current grim situation in American education. This study examined international data to see how the US education performed in the global competition, and how education in the US could be improved. On March 20, 2012, a ***Wall Street Journal*** report reminded us of the grim situation and the urgency of the education problem facing the US:

"Flaws in U.S. schools are increasingly causing a national-security risk, producing adults without the math, science and language skills necessary to ensure American leadership in the 21st century, warns a report issued Tuesday by the Council on Foreign Relations. Warning that 'the education crisis is a national security crisis,' the report says that too many schools are failing to adequately equip students for the work force, and that many have stopped teaching the sort of basic civics that prepare students for citizenship... The situation, it says, puts the country's 'future economic prosperity, global position, and physical safety at risk.'"

**Purpose Statement**

The purpose of this study was to use a newly available dataset collected from about 470,000 students in 17,145 schools in 65 countries in 2009 to determine the quantitative effects of class sizes on student performance, controlling for the economic, social and cultural status of students. In 2009, the PISA collected data from large samples of students from over 60 economies around the globe. These economies made up close to 90% the total size of the world economy (our analysis focuses on 40 economies because of data availability for PISA scores, class sizes and teacher pay). PISA recorded every student's standardized test scores and the student's economic, social and cultural status.

This study attempted to answer the following research questions: did students in countries with smaller classes (i.e., classes with higher pupil-teacher ratios) perform better academically? Did students in countries with higher teacher pay perform better academically? A systematic analysis of the dataset provided new insights in response to these important research questions. Policies based on such systematic scientific analysis would probably be more sound and effective than policies based on "conventional wisdom" without empirical evidence. This research project was thus not only important for parents and students, but also important for policy makers.

**Hypothesis Development**

Hypothesis 1: Class size had significant and negative effects on student learning outcomes. To test this hypothesis, this study measured the student learning outcomes by their standardized test scores in PISA (Programme for International Student Assessment). The dependent variables were three student performance measures in PISA: (1) reading literacy test score; (2) mathematical literacy test score; (3) scientific literacy test score.

The key independent variable was the class size. The control variables were explained in Chapter 3.

Hypothesis 2: Teacher pay had significant and positive effects on student learning outcomes. To achieve the best student learning outcome, it was essential for students to be taught by well-incentivized teachers. To test this hypothesis, the dependent variables were still the students' PISA scores, and the key independent variable was teacher's salary.

**Definitions of Terms**

The key dependent variables in this study were students' PISA scores in reading, mathematics and science, respectively. The key independent variables in this study were average class sizes and teacher's average salary.

- Average class sizes were the number of students in a typical class in the country.
- Teacher's pay was measured by the percentage of average salary for a teacher with 15 years of experience over per capita GDP. The advantage of using this relative income measure was that it effectively controlled for the differences in income levels between high-income and low-income countries.

The main control variables were as follows:

- The average time a student spent on regular lessons per week. It was expected that more time spent on regular lessons would boost student performance.
- Students' economic, social and cultural status, which was measured by the PISA Index of Economic, Social and Cultural Status. This index was created on the basis

of the following variables: the International Socio-Economic Index of Occupational Status (ISEl); the highest level of education of the student's parents, converted into years of schooling; the PISA index of family wealth; the PISA index of home educational resources; and the PISA index of possessions related to "classical" culture in the family home. For example, the PISA index of possessions related to "classical" culture in the family home measures how many books and computers are available in the family home.

**Assumption**

The main assumption in this research was that the relationship between class size, teacher pay and student performance was similar for countries in the sample. Even if this assumption were violated, as long as there were no systematic errors in the collection of PISA data, the empirical analysis in this study was still valid.

**Ethics**

Since this project used publicly available data collected from the website of PISA, there was no expectation of any ethical issues to be involved in it requiring the attention of the researcher.

In summary, it was very important to quantify the effects of class sizes and teacher pay on students' academic performance. A comprehensive analysis of a new sample of PISA surveys of students in 40 countries around the world in 2009 would shed light on this important topic. The results of this study provided new evidence to better understand the relationship between class size, teacher pay and student performance, and reinforce the theoretic foundation for relevant educational policies. It was hoped that this new scientific evidence could help policy makers better evaluate the benefits and costs of smaller classes and higher teacher pay and make better educational policies.

The remainder of this thesis was structured as follows: Chapter 2 reviewed the relevant literature. The research methodology was discussed in Chapter 3, and the research results were presented in Chapter 4. Chapter 5 summarized and concluded this thesis.

# Chapter 2

## Literature Review

Many researchers identified various determinants of student academic performance, such as class size, family background, teacher quality, educational expenditure, school autonomy, competition from private schools, and the composition of a school or classroom ("peer effects").

### Literature on Class Size

Previous researchers examined class size in a cross-country context but did not reached a consensus regarding its effect on students' academic performance. For example, Fuchs and Woessmann (2007) and Jurges and Schneider (2004) found no positive associations between reduced class sizes and student performance, while Barro and Lee (2001) found that lower class size (i.e., lower pupil-teacher ratio) in primary schools was associated with significantly higher student performance. Hanushek and Woessmann's (2011) international evidence showed that significant class-size effects were only present in systems with relatively low teacher quality. This result raised the cost-effectiveness question of whether student achievement would be best served by reducing class size or by improving the teacher quality. On a different but related issue, school size, instead of class size, was shown to have positive and significant effects on student performance in Jurges and Schneider (2004)'s study, i.e., students at larger schools seemed to perform better. They argued this was because larger schools tended to have more resources at the school level.

In addition to the above cross-country analysis, researchers conducted numerous within-country analyses. Some studies evaluated the effects of policy experiments.

Among these policy experiments, the most prominent one was the Project STAR (Student-Teacher Achievement Ratios) that had started in Tennessee in 1985. The project randomly assigned about 6,500 students in 329 classrooms in 79 schools to either regular classes (22-26 students) or small classes (13-17 students). Finn and Petrilli (1998) and Nye et al. (2000) reported students that had attended small classes obtained higher reading and mathematics scores, were more engaged in school, and had fewer disruptive or withdrawn behaviors compared to their counterparts in regular-sized classes.

Another important policy experiment was conducted in 1996 in California. According to Jepsen and Rivkin (2009, p223), it was the "most expensive state-level education reform in U.S. history." Specifically, California's class-size reduction program reduced K-3 class sizes throughout the state by roughly 10 students per class, at an annual cost exceeding one billion dollars. According to the estimation by Jepsen and Rivkin (2009), this reform did raise mathematics and reading achievement. However, an unintended consequence was that 25,000 new teaching positions needed to be filled in the first two years of the reform. Many of these positions were filled by teachers without certification or prior teaching experience. Their research highlighted the importance of considering class size and teacher quality at the same time. Rivkin et al. (2005) examined a longitudinal data set of test scores spanning grades 3 through 7 for 3 cohorts of Texas students in the mid-1990s and concluded that reductions in class sizes and better teacher quality were substitutes for each other.

Within-country studies were not only conducted in the US, but also in other countries. In the United Kingdom, Blatchford et al. (2002) followed a cohort of 9330 students in 368 classes in 220 schools over a 3-year period. They found that decreasing class size was related to increasing test scores, and the improvements were larger for the low achievers and those from lower socioeconomic backgrounds.

In Denmark, Heinesen (2010) examined French language test scores of over 7000 students and found highly significant and substantial positive effects of reducing class size. Such effects were larger for academically weaker students and for boys.

In Australia, Johnson et al. (2004) analyzed a data set based in Victoria generated from 1232 primary schools, 264 secondary schools and 44 primary and secondary schools. They concluded that (p.33), “we have been unable to find any evidence that class size is an important determinant of academic performance in primary or secondary schools.” Similar results were obtained when Lindahl (2001) examined 556 Swedish fifth graders to analyze the effects of class size on students’ academic performance. Similar conclusions were also reached when Fuller (1987) reviewed nine studies in 12 developing countries (Botswana, Thailand, India, Chile, Egypt, Kenya, Malaysia, Puerto Rico, Tanzania, Bolivia, and Argentina).

Current evidence about the effects of class size on students’ academic performance was inconclusive. Therefore, further research using updated PISA data was needed.

**Literature on Teacher Pay**

Figlio (1997) examined teacher-level data from the restricted-use version of the Schools and Staffing Survey (SASS), administered by the US Department of Education's National Center for Education Statistics, and found a significant relationship between teacher salaries and quality (measured by undergraduate college selectivity and subject matter expertise) within local teacher labor markets. Based on this evidence, Figlio (1997) concluded that public schools that paid higher salaries apparently attracted higher quality teachers within local teacher labor markets.

On the other hand, Eide, Goldhaber and Brewer (2004) examined the occupational choices of prospective teachers in the US and found that teachers were disproportionately drawn from the lower end of the academic proficiency distribution. This was in sharp contrast with the situation in the 1960s, when the most academically proficient college graduates were as likely to enter teaching as any other occupation. Eide, Goldhaber and Brewer (2004) argued that the low pay at US public schools made them harder and harder to attract the best and brightest college graduates.

Lakdawalla (2006) examined the long-run trend of wages of US teachers relative to other skilled workers with similar experience, and found that between 1940 and 1990, teachers' relative wages had fallen significantly. This long-run trend was still continuing today, which was quite alarming. This trend might have contributed to the grim "education crisis" described at the beginning of Chapter 1 of this thesis. This trend was jeopardizing the leadership position of the US in the world. For example, between 1995 and 2008, the United States slipped from ranking second in college graduation rates to 13th, according to the Organization for Economic Co-operation and Development (OECD), the Paris-based organization that develops and administers the PISA exam. Of 34 OECD countries, only 8 had a lower high school graduation rate.

Research on teacher pay focused almost exclusively on US teachers. Cross-country studies on this issue were quite sparse. This study intended to fill this void by providing evidence on this important issue with rigorous analysis of cross-country data.

**Literature on Other Determinants of Students' Academic Performance**

In addition to class size, researchers examined various factors that might account for the differences in student performance across countries, notably family background, teacher quality, school autonomy, competition from private schools, educational expenditure, and the composition of a school or classroom ("peer effects").

A common feature of the literature on this issue was that all studies examined large samples of students from multiple countries, which allowed for rich variations of country-level, school-level, class-level and student-level variables. Since this study also conducted cross-country research, the surveyed studies in the literature could provide useful guidance for this study.

For example, Woessmann (2003) studied the academic performance of 260,000 students from 39 countries, Fuchs and Woessmann (2007) examined the academic performance of 96,000 students from over 6,600 schools in 31 countries, Jurges and Schneider (2004) investigated the academic performance of 130,791 students from 23 countries. Another common feature of these studies was that they all used international standardized test scores to measure students' academic performance, though the last study (Lee and Barro, 2001) also used some additional measures of the educational system such as school repetition and dropout rates to measure students' performance. Three international standardized tests that were most commonly used in this literature were PISA (Programme for International Student Assessment), TIMSS (Trends in International Mathematics and Science Study), and IEA (International Association for the Evaluation of Educational Achievement). PISA had three major components, math, science and reading. It is the most comprehensive international standardized test. TIMSS focused on math and science, while IEA focused on math alone. Woessmann (2003), Jurges and Schneider (2004) used TIMSS data, Fuchs and Woessmann (2007) used PISA data, Zimmer and Toma (2000), Heyneman and Loxley (1983), and Lee and Barro (2001) used IEA data.

Researchers found that family background had significant effects on student's academic performance. Woessmann (2003), Hanushek and Woessmann (2011), Lee and Barro (2001) and Jurges and Schneider (2004) all showed that family background significantly affected student performance. For example, Woessmann (2003) showed in a cross-country study that the educational level achieved by the students' parents was strongly and positively correlated with the students' educational performance. He also found that the effect of the variable 'books at home,' which represented the educational and social environment in the students' home, was even stronger.

Examining data from England, Hanushek and Woessmann (2011) showed that the difference in educational achievement between children of families with more than two bookcases of books and children of families with only very few books at home (the two extremes of the five available categories) is 1.15 standard deviations, or more than three times what students on average learn during a whole school year.

Barro and Lee's (2001) result also suggested that parents' income and education had an important positive effect on the children's test scores. Jurges and Schneider (2004, p.367) concluded that "the strongest determinants of individual student performance are a student's social background variables such as the parent's formal education, immigrant status, the number of books at home, whether there is a personal computer at home, or parental attitudes towards doing well in math." (They studied the TIMSS dataset, which only included math test scores.)

In addition to family background, researchers also found that teacher quality had significant effects on student performance. Heymann and Loxley (1983, p.1162) found teacher quality had significant and positive effects on student performance, especially in low-income countries. They concluded, "The predominant influence on student learning is the quality of the schools and teachers to which children are exposed." Jurges and Schneider (2004) found that teachers' training and education seemed to benefit weak students more than strong students. They also found that if a teacher gave homework more frequently, students' academic performance was significantly better.

Researchers found that school autonomy had significant effects on student performance as well (Woessmann, 2003; Hanushek and Woessmann, 2011; Fuchs and Woessmann ,2007; and Jurges and Schneider,2004).

Woessmann (2003) showed that students in schools with autonomy in deciding on the hiring of teachers performed statistically significantly better in mathematics and science, as did students in schools that could determine teacher salaries themselves. Jurges and Schneider (2004)'s results were consistent with those in Woessmann (2003). They showed that students in schools with autonomy in determining teacher salaries, hiring teachers, and formulating the school budget performed significantly better than students in schools without autonomy in any of the above three domains .

But these three domains were not the only domains in which school autonomy matters. Other autonomies, such as the autonomy of selecting teaching methods or textbooks, were also important. For example, Hanushek and Woessmann (2011) showed that students performed better if their teachers had both incentives and the possibility to select appropriate teaching methods. The effects of school autonomy also depended on other factors. For example, Fuchs and Woessmann (2007) found that autonomy was more positively associated with performance in education systems that had external exit exams. They concluded that external exams and school autonomy were complementary, in that external exit exams provided “quality control” while school autonomy allowed schools to make optimal decisions that worked best for each individual school.

Researchers found that on average, private schools students performed better than public school students. Woessmann (2003) , Hanushek and Woessmann (2011), and Fuchs and Woessmann (2007) all showed that private school students tended to perform better than public school students. Woessmann (2003) showed that students in countries with larger shares of enrollment in privately managed educational institutions scored statistically higher in both mathematics and science. Fuchs and Woessmann (2007) showed that students in privately managed schools performed statistically better than students in publicly managed schools, after controlling for a large set of student background differences.

In terms of the effects of the composition of a school or classroom on academic performance of students, Zimmer and Toma (2000) showed that “peer effects” were significant determinants of educational achievement, and such effects appeared to influence low-ability students more than high-ability students. Specifically, they argued that the composition of a school or classroom—that was, the characteristics of the students themselves—affected the educational attainment of an individual student. This influence of the students in a classroom was often referred to as a “peer effect.”

Researchers found that educational expenditure had mixed effects on student performance. Hanushek and Woessmann's (2011) international evidence on the role of educational expenditure in educational production provided little confidence that quantitative measures of expenditure and class size were major drivers of student achievement, across and within countries. On the other hand, Woessmann (2003) showed that students in countries with a higher share of educational expenditure spent on private institutions performed statistically significantly better in mathematics, again attesting to the effectiveness of private educational institutions.

In summary, researchers found that good family background, well-trained teachers, high school autonomy, competition from private schools and a favorable composition of a school or class ("peer effects") all had positive and significant effects on students' academic performance, while the effects of class size and educational expenditure seemed to be controversial.

To investigate the effects of class size and educational expenditure on students' academic performance, this study used rigorous analytical methodology. Chapter 3 discussed the specific methodology in this study. Specifically, Chapter 3 discussed the data collection and analysis procedures.

# Chapter 3

# Methodology

## Data Collection

The data for this study came from PISA (Programmed for International Student Assessment) 2009. In 2009, around 470,000 15-year-old students from 17,145 schools in 65 countries completed the assessment, representing about 26 million 15-year-olds in the schools of the 65 participating countries and economies (OECD, **2010).** These 65 countries and economies made up close to 90% of the world economy. However, not all variables were available for all 65 countries. For example, some countries had teacher's salary data but not class size data, while some countries had class size data but not teacher's salary data. There were 40 countries with all variables available for the purpose of this study, so the sample size for this study was 40.

Each participating student spent two hours completing pencil-and-paper tasks in reading, mathematics and science. The assessment included tasks requiring students to construct their own answers as well as multiple-choice questions. The latter were typically organized in units based on a written passage or graphic, much like the kind of texts or figures that students might encounter in real life. In addition to the tests in reading, mathematics and science, students also answered a questionnaire that took about 30 minutes to complete. This questionnaire focused on their personal background, their learning habits, their attitudes towards reading, and their engagement and motivation. Meanwhile, the school principals completed a questionnaire about their school that included demographic characteristics and an assessment of the quality of the learning environment at school.

**Procedures and Data Analysis**

To examine the effects of class size and teacher pay on students' academic performance, this study tested the following two main hypotheses:

Hypothesis 1: Class size had significant and negative effects on student learning outcomes, after controlling for other influential factors such as the student's social economic status and the time students spent on regular lessons.

Hypothesis 2: Teacher pay had significant and positive effects on student learning outcomes, after controlling for other potentially influential variables.

A quantitative methodology was employed in this study to estimate the effects of class size and teacher pay on student performance in reading, mathematics and science. To test the first hypothesis, regressions of students' PISA test scores on class size were conducted. To test the second hypothesis, regressions of students' PISA test scores on teacher pay were conducted.

Since the PISA 2009 scores are based on a 600-point scale, while the maximum class size in the sample is about 40, to test Hypothesis 1, it was more meaningful to examine how much ***percentage*** increase or decrease in PISA 2009 scores would result from a ***percentage*** increase or decrease in the class size. To achieve this purpose, the natural logarithm transformation of PISA scores and class sizes was used. The following simple regression equation was estimated with the Ordinary Least Squares (OLS) method:

$$\ln(PISA\ score) = constant + beta1 * \ln(class\ size) + error \quad (1)$$

To see the logic behind the above regression equation, we derived the marginal effect of a change in the independent variable (ln(class size)) on the dependent variable as follows:

$$\frac{d\ln(PISA)}{d\ln(\text{Class size})} = \frac{d(\text{PISA Score})/\text{PISA Score}}{d(\text{Class size})/\text{Class size}} = \frac{\text{Percentage change in PISA Score}}{\text{Percentage change in class size}}$$

A t-test on the coefficient betal was conducted. The null hypothesis was: betal equals 0, while the alternative hypothesis was: betal does not equal 0. If the p-value from the t-test was less than 0.1 but greater than 0.05, then we rejected the null hypothesis at 10% significance level. If the p-value from the t-test was less than 0.05 but greater than 0.01, then we rejected the null hypothesis at 5% significance level. If the p-value from the t-test was less than 0.01, then we rejected the null hypothesis at 1% significant level. In all these three cases, we concluded that class size did have statistically significant effects on student performance. However, if we found the p-value from the t-test to be greater than 0.1, then we were not able to reject the null hypothesis, and we had to conclude that we were unable to find evidence that class size had statistically significant effects on student performance.

To test hypothesis 2, the following simple regression equation was estimated with the OLS method:

***ln(PISA score)=constant+beta2 *Teacher Pay+error term*** (2)

Recall that "Teacher Pay" was measured by the ratio of average upper middle school teacher's salary over per capita GDP, beta2 could be interpreted as the percentage change in PISA score if we increased "Teacher Pay" by one percentage point. This was because the marginal effect of "Teacher Pay" on "ln(PISA score)" could be computed as follows:

$$\frac{d\ln(PISA)}{d(\text{Teacher Pay})} = \frac{d(PISA)/PISA}{d(\text{Teacher Pay})} = \frac{\text{Percentage change in PISA score}}{\text{change in Teacher Pay}}$$

Similar to equation (1), we could conduct a t-test for beta2 to see if "Teacher Pay" had statistically significant effects on student performance.

A serious limitation of the simple regression analysis was that it omitted some important determinants of student performance, as discussed in Chapter 2. To conduct a rigorous analysis of the effects of class sizes on student performance, multivariate regressions were conducted. Specifically, the multivariate regression equation to be estimated with OLS method was as follows:

***ln(PISA score)=constant+betal*ln(class size)+beta2*Teacher Pay+control variables+error term*** (3)

In equation (3), "control variables" included the time students spent on regular classes per week (in minute), the social-economic status of students, and the ratio of educational expenditure per student over per capita GDP.

The above analytical methodology was employed for a systematic research on how class size and teacher pay affected students' academic performance. The results of the empirical analysis were reported in Chapter 4.

# Chapter 4

## Research and Discussion

### Summary Statistics

Students' academic performance was measured by their scores in mathematics, science and reading, respectively. Table 1 listed the average scores for students in each country in our sample of 40 countries in PISA 2009. Students from Shanghai (China) scored highest in all three subjects: math, science and reading, while students from Kyrgyzstan scored lowest among the 40 countries in the sample.

Table 1. Average PISA Scores (Ranking by math score, from highest to lowest)

| country | Math | science | reading |
|---|---|---|---|
| Shanghai-China | 600 | 575 | 555.8 |
| Singapore | 562 | 542 | 525.9 |
| Hong Kong-China | 555 | 549 | 533.2 |
| Korea | 546 | 538 | 539.3 |
| Chinese Taipei | 543 | 520 | 495.2 |
| Finland | 541 | 554 | 535.9 |
| Switzerland | 534 | 517 | 500.5 |
| Japan | 529 | 539 | 519.9 |
| Netherlands | 526 | 522 | 508.4 |
| Macao-China | 525 | 511 | 486.6 |
| New Zealand | 519 | 532 | 520.9 |
| Australia | 514 | 527 | 514.9 |
| Germany | 513 | 520 | 497.3 |
| Estonia | 512 | 528 | 501 |
| Iceland | 507 | 496 | 500.3 |
| Denmark | 503 | 499 | 494.9 |
| Slovenia | 501 | 512 | 483.1 |
| Norway | 498 | 500 | 503.2 |
| France | 497 | 498 | 495.6 |
| Austria | 496 | 494 | 470.3 |

| country | math | science | reading |
|---|---|---|---|
| Poland | 495 | 508 | 500.5 |
| Sweden | 494 | 495 | 497.4 |
| Czech Republic | 493 | 500 | 478.2 |
| Hungary | 490 | 503 | 494.2 |
| Luxembourg | 489 | 484 | 472.2 |
| United States | 487 | 502 | 499.8 |
| Portugal | 487 | 493 | 489.3 |
| Ireland | 487 | 508 | 495.6 |
| Spain | 483 | 488 | 481 |
| Italy | 483 | 489 | 486.1 |
| Greece | 466 | 470 | 482.8 |
| Croatia | 460 | 486 | 475.7 |
| Israel | 447 | 455 | 474 |
| Bulgaria | 428 | 439 | 429.1 |
| Thailand | 419 | 425 | 421.4 |
| Montenegro | 403 | 401 | 407.5 |
| Colombia | 381 | 402 | 413.2 |
| Qatar | 368 | 379 | 371.7 |
| Peru | 365 | 369 | 369.7 |
| Kyrgyzstan | 331 | 330 | 314 |

Among these 40 countries, students in the United States ranked 26 in math score, 19 in science score, and 15 in reading score. This confirmed the grim situation of the "education crisis" in the US, as the beginning of this thesis indicates.

Table 2 listed the average class sizes in out sample of 40 countries in PISA 2009. The smallest class size was 18.6 (Switzerland), while the largest class size was Chinese Taipei (39.5).

**Table 2. Average Class Sizes (Ranking from Smallest to Largest)**

| Country | class size | | country | class size |
|---|---|---|---|---|
| Switzerland | **18.6** | | New Zealand | 24.2 |
| Iceland | 18.7 | | United States | 24.4 |
| Finland | 19.2 | | Germany | 24.8 |
| Denmark | 19.4 | | Qatar | 25.9 |
| Austria | 20.8 | | Croatia | 26.2 |
| Italy | 20.9 | | France | 26.9 |
| Sweden | 21 | | Montenegro | 28.1 |
| Luxembourg | 21 | | Slovenia | 28.2 |
| Spain | 21.8 | | Israel | 28.5 |
| Kyrgyzstan | 22.1 | | Hungary | 28.5 |
| Portugal | 22.3 | | Peru | 28.9 |
| Bulgaria | 22.4 | | Singapore | 34.9 |
| Poland | 22.5 | | Colombia | 35.1 |
| Estonia | 22.5 | | Hong Kong-China | 35.6 |
| Greece | **22.6** | | Korea | 35.9 |
| Ireland | 22.7 | | Japan | 37.1 |
| Australia | 22.9 | | Thailand | 37.7 |
| Norway | 23.4 | | Macao-China | 38.4 |
| Netherlands | 23.7 | | Shanghai-China | 39 |
| Czech Republic | 24 | | Chinese Taipei | 39.5 |

Table 3 (in page 27) listed the average upper middle school teacher pay/per capita GDP. Crotia paid their teachers least generously, while Hong Kong (China) paid their teachers most generously.

Table 4 reported the descriptive statistics for the key dependent and independent variables. There were sufficient variations in all these variables to make subsequent statistical analysis meaningful.

**Table 3. Average Upper School Teacher Pay/Per Capita GDP**

| Country | Teacher Pay/GDP | | Country | Teacher Pay/GDP |
|---|---|---|---|---|
| Croatia | 0.38 | | Macao-China | 1.23 |
| Qatar | 0.5 | | Finland | 1.26 |
| Estonia | 0.61 | | Ireland | 1.26 |
| Norway | 0.69 | | Australia | 1.27 |
| Israel | 0.82 | | Montenegro | 1.34 |
| Iceland | 0.87 | | Denmark | 1.4 |
| Hungary | 0.94 | | New Zealand | 1.42 |
| Czech Republic | 0.97 | | Japan | 1.44 |
| Peru | 0.97 | | Colombia | 1.46 |
| Sweden | 0.98 | | Chinese Taipei | 1.55 |
| Bulgaria | **1** | | Portugal | 1.55 |
| United States | 1.01 | | Spain | 1.56 |
| Kyrgyzstan | 1.02 | | Netherlands | **1.66** |
| France | 1.05 | | Singapore | 1.67 |
| Poland | 1.1 | | Shanghai-China | 1.75 |
| Italy | 1.13 | | Switzerland | 1.8 |
| Greece | 1.13 | | Germany | 1.82 |
| Austria | 1.13 | | Korea | 2.01 |
| Luxembourg | 1.18 | | Thailand | 2.19 |
| Slovenia | 1.18 | | Hong Kong-China | 2.34 |

**Table 4. Descriptive Statistics for Key Dependent Variables and Independent Variables in the Regressions**

| Variable | Mean | First Quartile | Median | Third Quartile | Min | Max | Standard Deviation |
|---|---|---|---|---|---|---|---|
| Reading Score | 480.89 | 474.85 | 495.05 | 502.1 | 314 | 555.8 | 49.24 |
| Math Score | 486.93 | 474.5 | 495.5 | 522 | 331 | 600 | 56.94 |
| Science Score | 489.98 | 485 | 500 | 521 | 330 | 575 | 52.96 |
| Teacher pay/GDP | 1.27 | 0.99 | 1.21 | 1.55 | 0.38 | 2.34 | 0.43 |
| Class Size | 26.51 | **22.2** | 24.1 | 28.7 | 18.6 | 39.5 | 6.36 |

Univariate Analysis

The univariate analysis of the sample showed that while teacher pay was positively and significantly correlated with student performance, class size was not significantly correlated with student performance. To make it easier for a graphical presentation of the above relationships, this study abbreviated country names, as shown in Table 5.

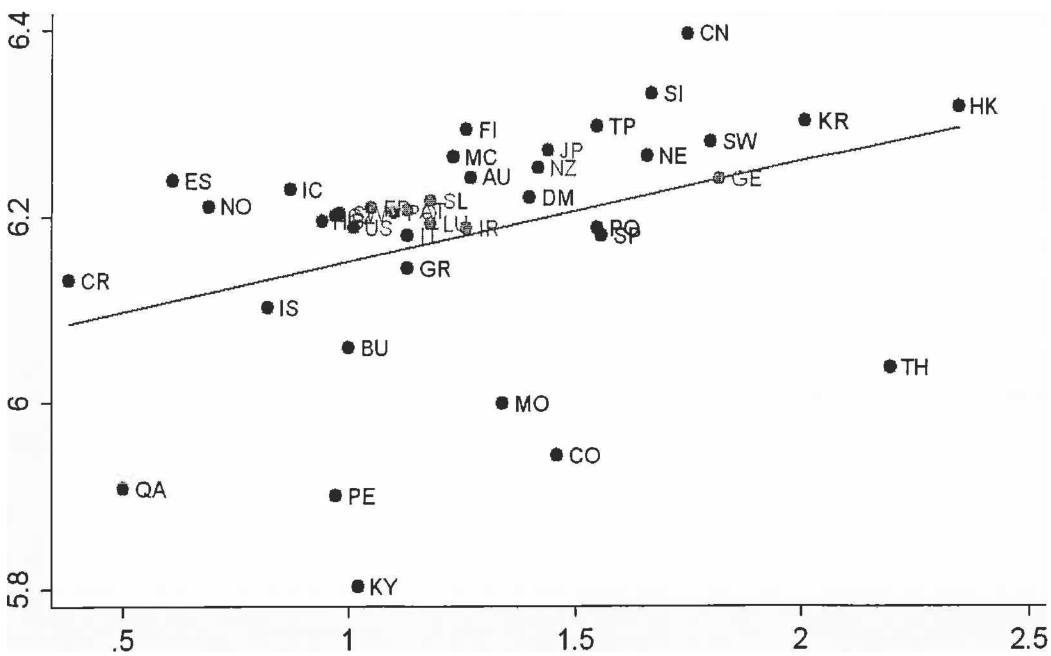
**Table 5. The Relationship between Abbreviated Country Name and Full Country Name**

| Abbreviated Name | Full Name | | Abbreviated Name | Full Name |
|---|---|---|---|---|
| AU | Australia | | KR | Korea |
| AT | Austria | | KY | Kyrgyzstan |
| BU | Bulgaria | | LU | Luxembourg |
| TP | Chinese Taipei | | MC | Macao-China |
| CO | Colombia | | MO | Montenegro |
| CR | Croatia | | NE | Netherlands |
| CZ | Czech Republic | | NZ | New Zealand |
| DM | Denmark | | NO | Norway |
| ES | Estonia | | PE | Peru |
| FI | Finland | | PO | Poland |
| FR | France | | PO | Portugal |
| GE | Germany | | QA | Qatar |
| GR | Greece | | CN | Shanghai-China |
| HK | Hong Kong-China | | SI | Singapore |
| HG | Hungary | | SL | Slovenia |
| IC | Iceland | | SP | Spain |
| IR | Ireland | | SW | Sweden |
| IS | Israel | | SW | Switzerland |
| IT | Italy | | TH | Thailand |
| JP | Japan | | US | United States |

In Figure 1, the horizon axis was average teacher pay for a teacher with 15 years of experience over per capita GDP in the country, while the vertical axis was the natural logarithm of the average math score of students in the country. Clearly, there was a positive relationship between teacher pay and students' math score. In this figure, there was a clear outlier: "Kyrgyzstan" (abbreviated as "KY"). In subsequent regression analysis, this outlier was removed from the sample to eliminate its undue influence on the regression results.

**Figure 1. Teacher Pay and Students' Average Math Score**

**Note: Teacher Pay was on the horizontal axis and the natural logarithm of Students' Math Score was on the vertical axis**




Similar positive correlations between teacher pay and students' science score, and between teacher pay and students' reading score, were also confirmed in Figure 2 and Figure 3, respectively.

In Figure 2, the horizon axis was average teacher pay for a teacher of 15 years of experience over per capita GDP in the country, while the vertical axis was the natural logarithm of the average science score of students in the country. In Figure 3, the horizon axis was average teacher pay for a teacher of 15 years of experience over per capita GDP in the country, while the vertical axis was the natural logarithm of the average reading score of students in the country. Similar to Figure 1, “Kyrgyzstan” (“KY”) was a clear outlier.

**Figure 2. Teacher Pay and Students’ Average Science Score**

**Note: Teacher Pay was on the horizontal axis and the natural logarithm of Students’ Science Score was on the vertical axis**

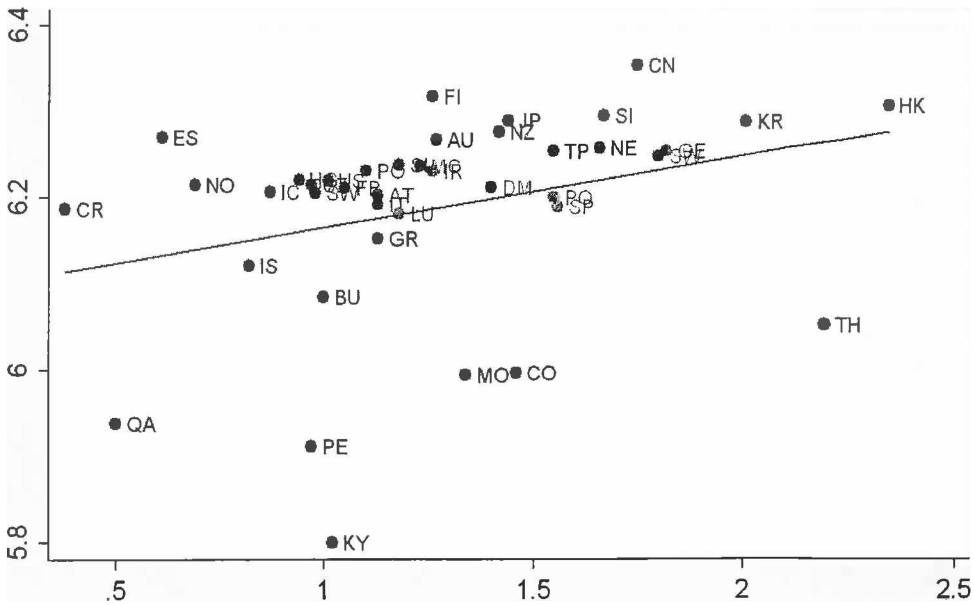



**Figure 3. Teacher Pay and Students’ Average Reading Score**

**Note: Teacher Pay was on the horizontal axis and the natural logarithm of Students’ Reading Score was on the vertical axis**

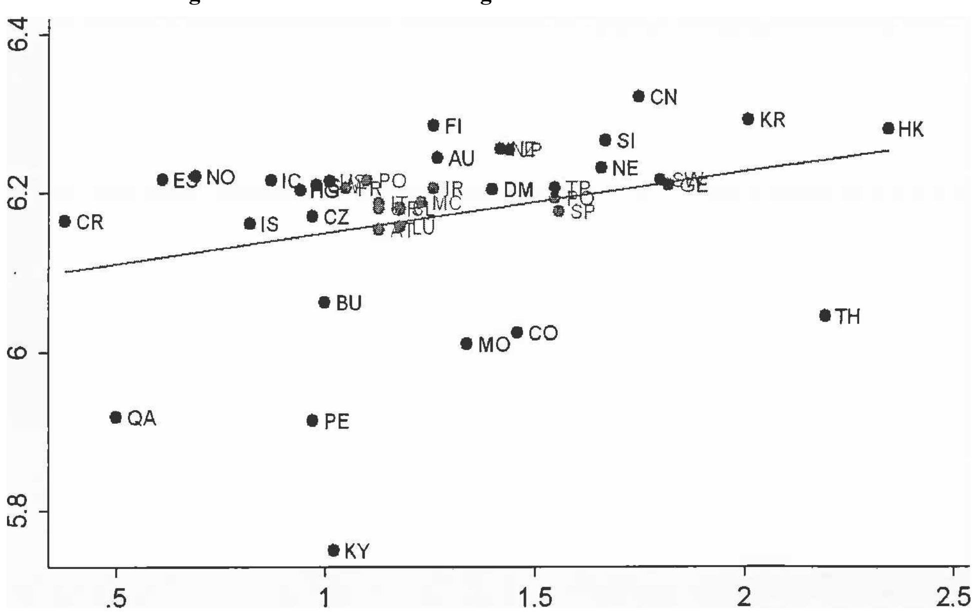


Figure 4 reported the relationship between the natural logarithm of class size and the natural logarithm of students' average math score. The horizontal axis was the natural logarithm of class size, and the vertical axis was the natural logarithm of students' average math scores. Clearly, "Kyrgyzstan" was an outlier here. It seemed that there was no identifiable relationship between class size and students' average math score.

**Figure 4. Class Size and Students' Average Math Score**

**Note: The natural logarithm of class size was on the horizontal axis and the natural logarithm of Students' Math Score was on the vertical axis**

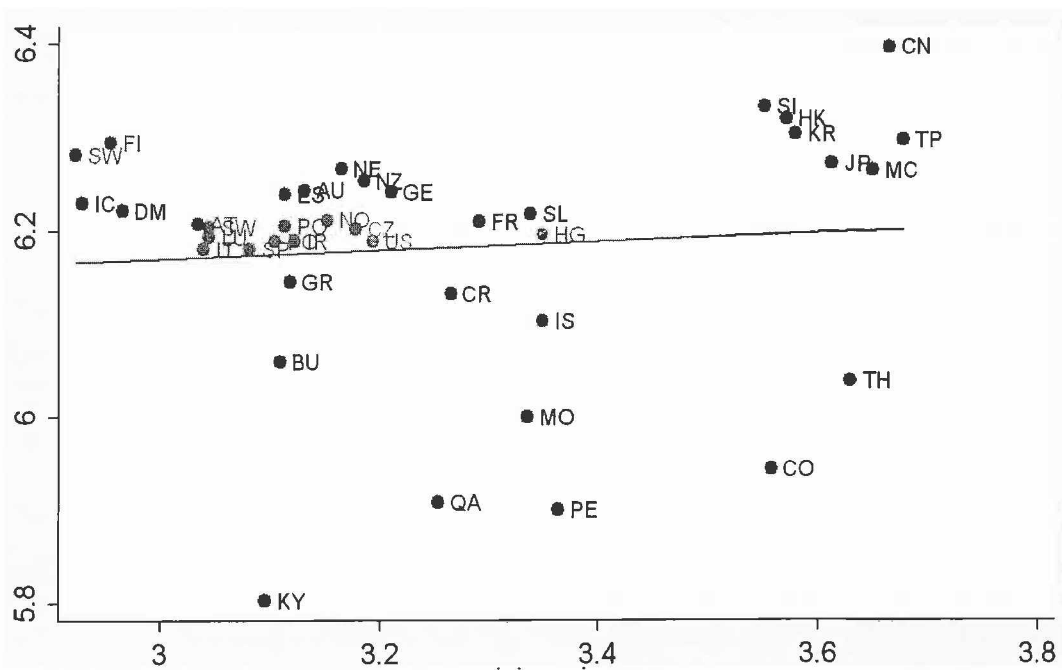



Figure 5 and Figure 6 reported the relationship between the natural logarithm of class size and the natural logarithm of students' average science and reading scores, respectively. The horizontal axis was the natural logarithm of class size, and the vertical axis was the natural logarithm of students' average score. Similar to Figure 4, it seemed that there was no identifiable relationship between class size and students' average science or reading score.

**Figure 5. Class Size and Students' Average Science Score**

**Note: The natural logarithm of class size was on the horizontal axis and the natural logarithm of Students' Science Score was on the vertical axis**

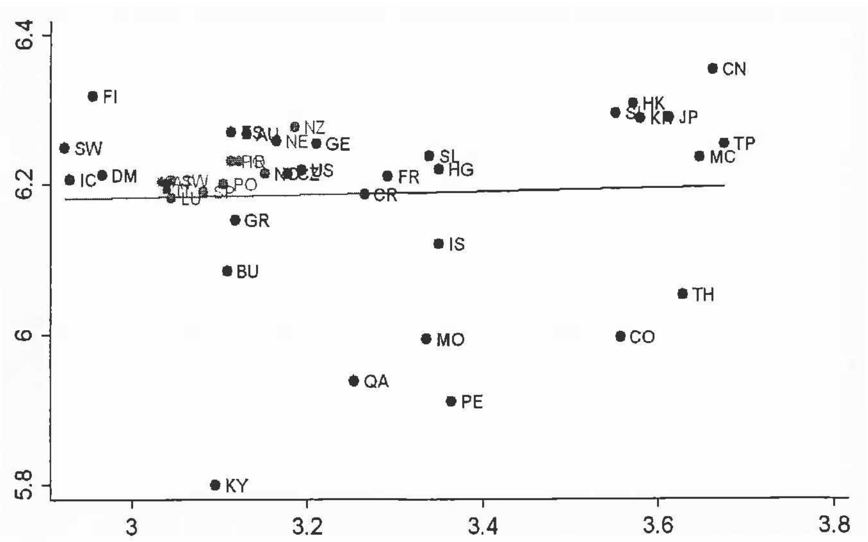
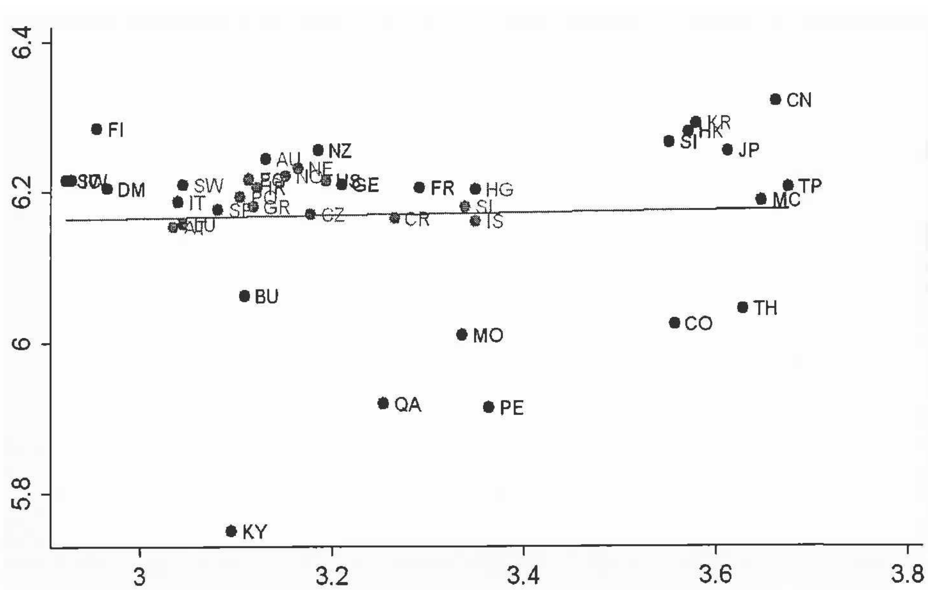



**Figure 6. Class Size and Students' Average Reading Score**

**Note: The natural logarithm of class size was on the horizontal axis and the natural logarithm of Students' Reading Score was on the vertical axis**



**Regression Analysis**

To quantify the effects of teacher pay and class size, regression analysis was conducted. Specifically, simple regressions of student performance on teacher pay and class size, respectively, was first conducted. Then multivariate regressions of student performance on teacher pay and class size, together with control variables, were conducted. In all the regressions, because the outlier "Kyrgyzstan" was dropped out, the sample size became 39.

Table 6 in page 34 reported the simple regression results. The p-values were reported in the parentheses. ***, ** and * denoted statistical significance at 1%, 5% and 10%, respectively. The results showed that while teacher pay had positive and significant effects on student performance, class size had no statistically significant effects on student performance. Furthermore, the explanatory power (R-squared) in models 4 to 6 (on the effects of class size on student performance) was close to 0. The marginal effects of teacher pay on student performance were large. For example, a one-standard deviation increase in teacher pay (0.43) would lead to about 4.1% (0.096*0.43) increase in students' average math score, about 3.5% increase in students' average science score, and about 3.1% increase in students average reading score. In contrast, the marginal effects of class size on student performance were either negative or positive.

Table 7 in page 34 reported the multivariate regression results. The first three models did not include control variables (the natural logarithm of time students spent on regular lessons per week, and PISA economic, social and cultural status). In all models, teacher pay was positively and significantly correlated with student performance, while class size did not have statistically significant effects on student performance. In models 4-6, the p-value for "teacher pay" was close to 0, indicating that teacher pay had highly

**Table 6. Teacher pay, class size and student performance:**

**simple regression results**

| | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 | Model 6 |
|---|---|---|---|---|---|---|
| | ln(math) | ln(science) | ln(reading) | ln(math) | ln(science) | ln(reading) |
| teacher pay | 0.096** | 0.071* | 0.063* | | | |
| | [0.018] | [0.058] | [0.062] | | | |
| ln(class size) | | | | 0.016 | -0.007 | -0.015 |
| | | | | [0.841] | [0.919] | [0.816] |
| constant | 6.068*** | 6.108*** | 6.100*** | 6.138*** | 6.222*** | 6.231*** |
| | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] |
| Observations | 39 | 39 | 39 | 39 | 39 | 39 |
| R-squared | 0.141 | 0.094 | 0.091 | 0.001 | 0.000 | 0.001 |

**Table 7. Teacher pay, class size and student performance:**

**multivariate regression results**

| | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 | Model 6 |
|---|---|---|---|---|---|---|
| | ln(math) | ln(science) | ln(reading) | ln(math) | ln(science) | ln(reading) |
| teacher pay | 0.110** | 0.086** | 0.079** | 0.145*** | 0.117*** | 0.107*** |
| | [0.014] | [0.036] | [0.033] | [0.000] | [0.002] | [0.001] |
| ln(class size) | -0.068 | -0.074 | -0.076 | 0.118 | 0.098 | 0.073 |
| | [0.409] | [0.338] | [0.273] | [0.172] | [0.234] | [0.318] |
| ln(time) | | | | 0.086 | 0.054 | 0.072 |
| | | | | [0.216] | [0.416] | [0.226] |
| status | | | | 0.172*** | 0.154*** | 0.139*** |
| | | | | [0.000] | [0.000] | [0.000] |
| constant | 6.272*** | 6.327*** | 6.328*** | 5.181*** | 5.459*** | 5.435*** |
| | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] | [0.000] |
| Observations | 39 | 39 | 39 | 39 | 39 | 39 |
| R-squared | 0.157 | 0.117 | 0.121 | 0.460 | 0.399 | 0.415 |

significant and positive effects on student performance. These effects were really large. For example, a one-standard deviation increase in teacher pay (0.43) would lead to about 6.2% increase in students' average math score, about 5% increase in students' average science score, and about 4.6% increase in students' average reading score. This translated into an improvement of the ranking of US students' math scores from No. 25 (487 points) to No. 12 (517 points) among the 40 countries in the sample. This also improved the ranking of US students' science scores from No. 19 (502 points) to No. 10 (527.1) among the 40 countries in the sample. Similarly, the ranking of US students' reading scores rose from No. 15 (499.8) to No. 7(522.8) among the 40 countries in the sample. In terms of control variables, in models 4-6, while the time spent on regular lessons had no statistically significant effects on student performance, students' economic, social and cultural status had positive and significant effects on average student performance. Models 4-6 also had good explanatory power. Their R-squared values were at least 0.4, which meant these models could explain at least 40% of the variations in students' average PISA scores in the country.

**Summary**

The results in this section showed that teacher pay had positive and significant effects on students' academic performance, while class size did not have significant effects. These results implied that given limited financial resources available at the government, it would be more effective to improve students' academic performance if those financial resources were used to increase teacher pay instead of reducing class size.

Chapter 5 summarized the whole thesis, discussed policy implications, and indicated future research directions.

# Chapter 5.

## Summary and Conclusions

The latest data from PISA (Program on International Student Assessment) showed that the academic performance of high school students in the United States did not rank high in the world. For example, among the 40 countries in the sample of this study, students in the United States ranked 26 in math score, 19 in science score, and 15 in reading score. As U. S Education Secretary Arne Duncan commented, "This is an absolutely wake-up call for America. The results are extraordinarily challenging to us and we have to deal with the brutal truth. We have to get much more serious about investing in education. ..we have to invest in reform, not in the status quo."

While it was probably clear to everyone that we should get serious about investing in education reform, it was probably unclear how we should do this—especially, given the current woeful government budget situation. Should the government invest its limited financial resources on the increase of teacher pay, or the reduction of class size? This was an important question that called for rigorous research. The answer to this question might help formulate effective education policies that could enhance the quality of education and boost the competitiveness of the US in the intense global competition.

This study carefully analyzed data collected through PISA from 40 countries through OLS regressions, and discovered that while teacher pay had positive, statistically significant and large effects on students' academic performance, class size had no significant effects. To the best of the author's knowledge, this was the first study that used the latest PISA data to examine the relationship between teacher pay and students' academic performance in a cross-country setting and made original contributions to the study on the determinants of students' academic performance.

Based on this evidence, this study recommended that the government should spend its limited financial resources on raising teacher pay. This would attract the best and brightest people to become teachers, and significantly improve the teacher quality in the US. As teacher quality was shown to be the most important determinant of students' academic performance, an increase in teacher quality would be expected to dramatically improve students' academic performance. Indeed, Table 3 in Chapter 4 showed that United States ranked 28 among 40 countries in terms of the ratio of the upper middle school teachers' average pay over per capita GDP. This disappointing ranking might partially explain why the academic performance of students in the United States lagged behind the academic performance of students from quite a few countries.

But it might not be sufficient to just raise teacher pay across the board. Future research would be needed on how to design a better structure of teachers' compensation. For example, intuitively, it might be a good idea to include a fixed salary and a bonus in teachers' compensation, where the bonus should depend on the students' academic performance. However, empirical evidence was inconclusive about this "good idea" yet.

For example, to test whether it was truly a "good idea", Springer et al. (2010) conducted a three-year study titled "Project on Incentives in Teaching (POINT)" in the Metropolitan Nashville School System from 2006-2007 through 2008-2009, where middle school mathematics teachers voluntarily participated in a controlled experiment. The purpose of the experiment was to assess the effect of financial rewards for teachers whose students showed unusually large gains on standardized tests. This study showed that students of teachers randomly assigned to the treatment group (eligible for bonuses based on student performance) did ***not*** outperform students whose teachers were assigned to the control group (not eligible for bonuses based on student performance), thus casting doubt on the "pay for performance" idea and calls for further in-depth study on this issue.

On the other hand, Figlio and Kenny (2007) combined data from the National Education Longitudinal Survey on schools, students, and their families with a survey conducted in 2000 regarding the use of teacher incentives. They found that test scores were higher in schools that offered individual financial incentives for good performance. However, their study might not be as clear about the causal relationship between teacher's incentive-pay and student performance as the study of Springer et al. (2010). Figlio and Kenny (2007) admitted that "the association between teacher incentives and student performance could be due to better schools adopting teacher incentives or to teacher incentives eliciting more effort from teachers." Therefore, further research would be needed to figure out the ***causal*** relationship between teacher's incentive pay and student performance.

Regarding class size, the empirical research in this study showed that class size had no statistically significant relationship with student performance. This was consistent with the conclusion that Hanushek (1999) made from a survey of large number of studies: "there is little systematic gain from general reduction in class size"(p. 33). However, the lack of a robust correlation between class size and student performance might also be because some other important influential factors might have not been examined at the same time. For example, it may be possible that small class size would improve student performance only when teacher quality stayed constant or improved. Otherwise, the reduction of class size might be at the cost of worse teacher quality. For example, Imazeki (2003) found that, "in California, teacher quality, as measured by education, experience and credentials, fell dramatically in the wake of a state-wide class-size reduction policy." (p. 159)

The current study was a quantitative study. In the future, interviews should be conducted to discover more factors that might affect student performance. Especially, those qualitative studies would be expected to reveal some determinants of student performance that were ignored in previous studies. These studies would deepen our understanding of various determinants of student performance and help policy makers make better education policies. This would help restore the leadership position of the United States in the world and ensure that United States would remain highly competitive in the 21st century.

**References**

Allegretto, S. A., Corcoran, S. P. & Mishel, L. (2004). How does teacher pay compare? Methodological challenges and answers. Washington, D.C.: Economic Policy Institute.

Blatchford, P., Goldstein, H., Martin, C., & Browne, W. (2002). A study of class size effects in English school reception year classes. *British Educational Research Journal,* 29, 790-730.

Eide, E., Goldhaber, D., Brewer, D. (2004) The teacher labor market and teacher quality. *Oxford Review of Economic Policy,* 20, 230-244.

Figlio, D. (1997) Teacher salaries and teacher quality. *Economics Letters,* 55,267-271.

Figlio, D. N. & Kenny, L. W. (2007) Individual teacher incentives and student performance. *Journal of Public Economics,* 91, 901-914.

Finn, J. D. & Petrillii, P. (1998). Class size and student at risk: What is known? What is next? Washington DC: US Department of Education.

Fuchs, T., & Woesmann, L. (2007). What accounts for international differences in student performance? A re-examination using PISA data. Empirical Economics, 32, 433- 464.

Fuller, B. (1987). What school factors raise achievement in the third world? Review of Educational Research, 29, 809-818.

Halbach, A., Ehrler, K, Zahorik, J. & Molnar, A. (2001). Class size reduction: from promise to practice. *Educational Leadership,* 58, 32-35.

Hanushek, E. A. (1999) The Evidence on Class Size, in Susan E. Mayer and Paul Peterson (ed.), *Earning and Learning: How Schools Matter* (Washington, DC: Brookings Institution), pp. 131-168.

Hanushek, E. A., & Rivkin, S. G. (2007) Pay, Working conditions, and teacher quality. *The Future of Children,* Vol. 17, No. 1.

Hanushek, E. A., & Woesmann, L. (2011). The economics of international differences in educational achievement, in Handbook of Economics of Education, 89-200.

Heinesen, E. (2010). Estimating class-size effects using within-school variation in subject-specific classes. Economic Journal, 120, 737-760.

Heyneman, S. & Loxley, W. (1983). The effect of primary-school quality on academic achievement across twenty-nine high and low-income countries. American Journal of Sociology, 88, 1162-1194.

Imazeki, J. (2003). Class-Size reduction and teacher quality: Evidence from California. School Finance and Teacher Quality: Exploring the Connection, David Monk and Margaret Plecki,159-178.

Jepsen, C.& Rivkin, S. (2009) Class size reduction and student achievement. Journal of Human Resources, 44, 223-250.

Johnson, D., Jensen, B, Feeny, S.& Methakullawat, B. (2004). Multivariate analysis of performance of Victorian schools. Working paper, Melboum Institute of of Applied Economics and Social Research, Melbourne, Victoria, Australia.

Jurges, H. & Schneider, K. (2004). International differences in student achievement: an economic perspective. German Economic Review, 5, 357-380.

Krueger, A. (2003). Economic considerations and class size. Economic Journal, 113, 34-63.

Lakdawalla, D. (2006) The economics of teacher quality. Journal of Law and Economics, 49, 285-329.

Lee, J., & Barro, R. J. (2001). Schooling quality in a cross-section of countries. Economica, 68, 465-488.

Lindahl, M. (2001). Home versus school learning: A new approach to estimating the effect of class size on achievement. IZA Discussion Paper.

Nye, B., Hedges, L.V.,& Konstantopoulos, S. (2000). The effects of small classes on academic achievement: The results of the Tennessee class size experiment. American Educational Research Journal, 37, 123-151.

OECD (2010). *PISA 2009 Results: What Students Know and Can Do —Student Performance in Reading, Mathematics and Science (Volume I)* .

Rivkin, S., Hanushek, E. & Kain, J. (2005). Teachers, schools and academic achievement. Econometrica, 73, 417-458.

Springer, M. G., Ballou, D., Hamilton, L. S., Le, V., Lockwood, J. R., McCaffrey, D. F., Pepper, M., Stecher, B. M. (2010). Teacher pay for performance: Empirical evidence from the project on incentives in teaching. National Center on Performance Incentives Working Paper.

Woessmann, L., (2003). Schooling resources, educational institutions and student performance: The international evidence. Oxford Bulletin of Economics and Statistics, 65, 117-170.

Zimmer, R. & Toma, E. (2000). Peer effects in private and public schools across countries. Journal of Policy Analysis and Management, 19, 75-92.

## Appendix

STATA Code for Quantitative Analysis in the Thesis

```
cd C:\Thesis
*Prepare data for analysis
use classsize09, clear
sort country
drop if cs==.
save classsize09,replace
use reading09,clear
sort country
drop if reading==.
gen lreading=ln(reading)
save reading09, replace
use math09,clear
drop if math==.
sort country
gen lmaths=ln(math)
save math09 .replace
use science09,clear
drop if science==.
sort country
gen lscience=ln(science)
save science09,replace
```

```
^Combine class size data and reading, math and science score data

use classsize09,clear

gen lclasssize=ln(cs)

rename cs classsize

merge 1:1 country using reading09

drop _merge

merge 1:1 country using math09

drop _merge

merge 1:1 country using science09

keep country lclasssize heading lmaths lscience classsize reading math science

sort country

save sc09

*Simple regression analysis

use sc09,clear

reg heading lclasssize if
regexm(country, "Shanghai ")==0&regexm(country, "Azerbaijan")==0

est store reading

reg lmaths lclasssize if
regexm(country,"Shanghai")==0&regexm(country,"Azerbaijan")==0

est store maths

reg lscience lclasssize if
regexm(country,"Shanghai")==0&regexm(country,"Azerbaijan")==0

est store science

esttab reading maths science using resultl.csv, ar2 compress nogap star(* 0.1 ** 0.05 ***
0.01) b(%6.3f) brackets p

graph twoway (lfit heading lclasssize) (scatter heading lclasssize, mlabel(country))
```

```
graph twoway (lfit lmaths lclasssize) (scatter lmaths lclasssize, mlabel(country))

graph twoway (lfit lscience lclasssize) (scatter lscience lclasssize, mlabel(country))

*Merge teacher pay data with student test score data

use pay09, clear

drop if regexm(country,"Belgium")==l&regexm(country,"Fl.")==l

replace country="Belgium" if regexm(country,"Belgium")==l

replace country="United Kingdom" if regexm(country,"England")==l

sort country

drop if startingpay==""

merge 1:1 country using sc09

keep if _merge==3

drop _merge

destring pay 15yr,replace ignore (",")

gen Ipl5=ln(payl5yr)

graph twoway (lfit heading lpl5) (scatter heading lpl5, mlabel(country))

graph twoway (lfit lmath lpl5) (scatter lmath lpl5, mlabel(country))

graph twoway (lfit lscience lpl5) (scatter lscience lpl5, mlabel(country))

save oecdpay

*Sort students' socioeconomic status data and study time data

use status,clear

sort country

save status,replace

use time,clear

sort country

save time,replace
```

```
*Combine data and run simple and multivariate regressions

use sc09,clear

sort country

merge 1:1 country using status

merge 1:1 country using time

merge 1:1 country using oecdpay

destring lss,replace

*simple  regression

reg  lreading  lpl5

est  store  reading

reg  lmaths  lpl5

est  store  maths

reg  lscience  lpl5

est store science

esttab reading maths science using result2.csv, ar2 compress nogap star(* 0.1 ** 0.05 ***
0.01) b(%6.3f) brackets p

*multivariate regression

reg lreading lpl5 lclasssize status lss

est store reading

reg lmaths lpl5 lclasssize status lss

est store maths

reg lscience lpl5 lclasssize status lss

est store science

esttab reading maths science using result3.csv, ar2 compress nogap star(* 0.1 ** 0.05 ***
0.01) b(%6.3f) brackets p
```